Question answered: "what will happen?" ? Prospective/forward ? Retrospective / backwards ?? Ambidirectional cohort (Both retrospective and prospective) ? e.g. Framingham study of cardiovascular disease: Started in 1948, 6000 citizens participated, followed for 20 years (study in 1970 by Gordon and Kannel) ? Possible uses: ? Typical cohort study ? Outcome assessment (patient outcomes: economic, functional, satisfaction, QOL, ..) ? Historical Cohort studies/AKA Retrospective cohort: Relies on prospective records collected (If accurate) – still forward in time in the past The defining characteristic of all cohort studies is that they track people forward in time from exposure to outcome. Data collection may be prospective or retrospective. Ex. Contraceptives and DVT. Lancet 2002; 359: 57–61 Research in Reverse ?In a case–control study the study group is defined by the outcome (e.g. presence of a disease), not by exposure to a risk factor. ?The study starts with identification a group of cases (individuals with a particular health outcome) in a given population and a group of controls (individuals without the health outcome) ?The frequency of exposure to a potential risk factor is then compared between cases and controls. ?If the frequency of exposure is more common among cases than controls, the exposure may be a risk factor for the outcome under investigation. ?Designed to help determine if an exposure is associated with an outcome (i.e., disease or condition of interest). ?In theory, the case–control study can be described simply. ? First, identify the cases and the controls ?Then, look back in time to learn which subjects in each group had the exposure(s), ?Then compare the frequency of the exposure in the case group to the control group. ? It is always retrospective because it starts with an outcome then traces back to investigate exposures As with all epidemiological investigations the beginning of a casecontrol study should begin with the formulation of a clearly defined hypothesis. ? Case definition ?It should be clearly defined at the onset of the study to ensure that all cases included in the study are based on the same diagnostic criteria. ? Source of cases ?The source of cases needs to be clearly defined. ? The first step in the selection of cases is the formulation of case definition. ? usually based on a combination of signs and symptoms, physical and pathological examinations, and results of diagnostic tests. ? Case identification and enrolment. ? Typical sources for identifying cases are hospital or clinic patient rosters, death certificates, special surveys, and reporting systems such as cancer or birth defects registries. ? Investigators consider both accuracy and efficiency in selecting a particular source for case. ? The goal is to identify as many true cases of disease as quickly and cheaply as possible. ?Selection of cases ?Case–control studies may use incident or prevalent cases. ?Incident cases ?comprise cases newly diagnosed during a defined time period. ?The use of incident cases is preferential, as the recall of past exposure(s) may be more accurate among newly diagnosed cases. ?In addition, the temporal sequence of exposure and disease is easier to assess among incident cases. ? Selection of cases ? Case–control studies may use incident or prevalent cases. ? Prevalent cases ? Comprise individuals who have had the outcome under investigation for some time. ? The use of prevalent cases may give rise to recall bias as prevalent cases may be less likely to accurately report past exposures(s). ? The interpretation of results based on prevalent cases may prove more problematic, as it may be more difficult to ensure that reported events relate to a time before the development of disease rather than to the consequence of the disease process itself. ? E.g, individuals may modify their exposure following the onset of disease ?The controls provide an estimate of the

exposure rate that would be expected to occur in the cases if there was no association between the study disease and exposure. ? Individuals selected as controls should be ?free of the disease ?similar to the cases in regard to the possibility of having past exposure during the time period of risk. ? Controls must be sampled independently of exposure status. ?Exposed and unexposed controls should have the same probability of selection. ? In case-control studies where cases are hospital based, it is common to recruit controls from the hospital population. ?However, the choice of controls from a hospital setting should not include individuals with an outcome related to the exposure being studied. For example In a case-control study of the association between smoking and lung CA, the inclusion of controls being treated for a condition related to smoking (e.g. chronic bronchitis) may result in an underestimate of the strength of the association between the exposure (smoking) and outcome (lung CA) ?Case-control studies are often used for uncommon outcomes, ??investigators often have a limited number of cases. ?In this situation the statistical power of the study can be increased somewhat by enrolling more controls than cases. ? Ratios greater than 4:1 have little additional impact on power. ? However, if the data on controls is easily obtained, there is no reason to limit the number of controls. ?Can be 1:1, 1:2, 1:3, 1:4 ?The process of making a study group and a comparison group similar or identical with respect to their distribution of extraneous factors ?This can be done by ?by individual matching (e.g. for each case, choose a control of the same age and gender) or ?by frequency matching (e.g. if there are 25 male cases in the 30- 34 age-group then choose the same number of male controls for this age-group). ?Case-control studies are used to investigate the risk of disease in relation to a wide variety of exposures, ?lifestyle, occupation, environment, genes, diet, reproduction, and the use of medications. ?Sources available for obtaining exposure data include ?in-person and telephone interviews; ?self-administered questionnaires; ?preexisting medical, pharmacy, registry, employment, insurance, birth, death, and environmental records; and ?biological specimens. ? Susceptible to bias introduced due to poor study design or during the collection of exposure and outcome data. ? Differential reporting of exposure information between cases and controls based on their disease status. ? Cases and controls may recall past exposure differently (Recall Bias) ? Similarly, the recording of exposure information may vary depending on the investigator's knowledge of an individual's disease status (interviewer/observer bias). ? Temporal bias (reverse causality) may also occur. When trying to establish a link between exposure and outcome, it must be clear that the exposure occurred before the disease of interest ? The odds ratio (OR) is used in C-C studies to estimate the strength of the association between exposure and outcome. ? It is not possible to estimate the incidence of disease ? The results of a C-C study can be presented in a 2x2 table ? The OR is a measure of the odds of exposure in the cases, compared to the odds of exposure in the control group. ? Calculation of the OR from a hypothetical case-control study of smoking and cancer of the pancreas among 100 cases and 400 controls individuals with cancer of the pancreas (cases) are 4.5 times more likely to have smoked than those without the disease ? OR = 1 ? The odds of exposure among cases = the odds among the controls. ? Exposure does not appear to be a risk factor. ? OR1 ? Larger odds of exposure among the cases than among the controls. ? Cases have a higher odds of having been exposed in the past. ? Exposure appear to increase risk of disease. ? OR1 ? Smaller odds of exposure among cases as compared to controls, ??Controls have a higher odds of

having been exposed. ? This situation illustrates a protective factor; exposure appears to reduce risk of disease. Advantage ? Cost-effective relative to other studies. ? No long follow up period (as compared to cohort studies). ? Efficient for the study of diseases with long latency periods. ? Efficient for the study of rare diseases. ? Good for examining multiple exposures Disadvantage ? Selective recall (Recall bias) ? Mortality bias if prevalent rather than incident cases. ? Temporal relationship unclear. ? Difficulty in defining control group. ? Not optimal for rare exposures. ? A case-control study to assess the relation between stroke and cigarette smoking. 101 stroke patients were compared with 137 healthy controls. 71 of the stroke patients reported ever smoked compared to 36 of the controls reported ever smoked ? Is there is relation between stroke and cigarette smoking? Explain your answer. ? Construct a 2 x 2 table Stroke +ve Stroke -ve Total Ever smoked +ve 71 36 107 Ever smoked - ve 30 101 131 Total 101 137 238 ? Calculate the odds ratio (OR) OR= 71x101/36x30 = 7171/1080=6.6 ? Interpretation: stroke patients are 6.6 more probable to have ever smoked compared to controls ? A study where the cases and controls are selected from individuals within an established cohort study. ?It is thus said to be 'nested' within the cohort study. ?Cases of a disease that arise within the defined cohort during the follow up period are identified, then a specified number of matched controls who have not developed the disease are selected from the same cohort. ?Analysis is carried out in the same way as a normal case control, with calculation of odds ratios. ?Strengths of nested case-control studies ?Relatively cheap and easy to conduct. ?Data related to exposure and confounding have already been collected. ?Can utilize the baseline data on exposure and confounding collected before the onset of disease, which reduces the ?potential for recall bias ?uncertainty regarding the temporal sequence between exposure and disease onset. ? Case-control studies have been used in a variety of situations to evaluate possible causes of rare conditions. ? Classic examples include the investigation of cases of childhood leukaemia near the nuclear procession plant at Sellafield in Cumbria ?In a cross-sectional study, data are collected on the whole study population at a single point in time to examine the relationship between disease (or other health related state) and other variables of interest. ?They provide a snapshot of the frequency of a disease or other health related characteristics in a population at a given point in time. This methodology can be used to assess: ? Burden of illness and risk factors ? Service needs ? Hypothesis generation ? Estimate prevalence of disease at a single time point. ? Estimate prevalence of risk factors at a single time point. ? provide a "snapshot" of diseases and risk factors simultaneously in a defined population ??A cohort study is the best way to identify incidence and natural history of a disease, and can be used to examine multiple outcomes after a single exposure ?Also called follow-up or incidence studies ?Begin with sample "Healthy Cohort" (i.e., subjects without the outcome yet) ?Start with Exposure status, then compare subsequent disease experience in exposed vs. unexposed. ?Purpose: Estimate the incidence of disease or estimate association between exposure and disease ?The control group (unexposed) should be similar in all important respects to the exposed, with the exception of not having the exposure This type of observational study is the one that most closely resembles intervention studies, except that allocation of subjects to the exposure is ?This type of study can be done by going ahead in time from the present (prospective cohort study) or, alternatively, by going back in time to comprise the cohorts and following them up to the present (retrospective cohort study) ? Prospective studies: carried out from the

present time into the future, can be tailored to collect specific exposure data, but there may be a long wait for events to occur, particularly where the outcome of interest is associated with old age. ? They can be expensive and are prone to high dropout rates. ?Retrospective Studies (Historical Cohort): look at medical events from some time point in the past up to the present time. ?The advantage is that the information is available immediately. ?There may for such studies, however. The defining characteristic of all cohort studies is that they track people forward in time from exposure to outcome. ? Disadvantages ?The difficulty in tracing subjects ?The quality of recorded information. ? Selection of study groups ? The aim of a cohort study is to select study participants who are identical with the exception of their exposure status. ? At the beginning of the study; all study participants must be free of the outcome under investigation and have the potential to develop the outcome under investigation. ? If the exposure is common, the study population can be selected before classifying individuals ? If the exposure is rare, the study population may be chosen on the basis of exposure, to ensure sufficient exposed individuals are enrolled. ?Measuring exposure ? Levels of exposure (e.g. packs of cigarettes smoked per year) are measured for each individual at baseline at the beginning of study and assessed at intervals during the period of follow-up. ?When several exposures are being considered simultaneously, the non-exposed group should comprise all those with none of the risk factors under investigation. ? Exposure data may be obtained from a number of sources including medical or employment records, standardized questionnaires, interviews and by physical examination. ?Measuring outcome ?Outcome measures may be obtained from various sources including routine surveillance of cancer registry data, death certificates, medical records or directly from the participant. ?Note that the method used to ascertain outcome must be identical for both exposed and unexposed groups to avoid measurement bias. ??Information on outcomes should be collected by members of the study team who are blind to participants' exposure status, to reduce the risk of observer bias. ?Methods of follow up ?The follow up is a major challenge, and it may take many years for a sufficient proportion of participants to have reached an outcome. ?A great deal of cost and time is required to ensure adequate followup of cohort members and to update measures of exposures and confounders, as well as monitoring participants' health outcomes. ?Failure to collect outcome data for all members of the cohort will affect the validity of study results ? A major source of potential bias in cohort studies ? Cohort members may die, migrate, change jobs or refuse to continue to participate in the study. ? In addition losses to follow up may be related to the exposure, outcome or both. ? For example individuals who develop the outcome may be less likely to continue to participate in the study. ? The degree to which losses to follow up are correlated with either exposure or outcome can lead to serious bias in the measurement of effect of exposure and outcome ?Related to the degree of accuracy with which subjects have been classified with respect to their exposure or disease status. ? Differential misclassification, when one group of participants is more likely to have been misclassified than the other, can lead to an over- or underestimate of the effect between exposure and outcome ?Analysis of a cohort study uses the ratio of either the risk or rate of disease in the exposed cohort, compared with the rate or risk in the unexposed cohort. ?Relative Risk {Risk(Rate) Ratio} ?Comparing disease occurrence among exposed with disease occurrence among comparison group (usually unexposed) in a ratio measure. Risk in exposed Risk in unexposed ? Relative

Risk= ?The relative risk can be calculated as follows ?The RR of 15 indicates that the risk of cancer of the pancreas is 15 times higher among smokers than non-smokers ? If RR = 1: Risk in the exposed equal to risk in unexposed (no association) ? If RR1: Risk in exposed greater than risk in unexposed (positive association; possibly causal) ? If RR1: Risk in exposed less than risk in unexposed (negative association; possibly protective) Advantage ? Temporal relationship clear. ? Can study course of disease development. ?Good for rare exposures. ?Decreases potential for many biases since disease has not occurred at time of classification. Disadvantage ?Time-consuming, expensive. ?Loss to follow-up and missing data - ex; drop-outs, attrition, migration. ?Not suitable for rare diseases. A cohort study was conduct to evaluate the relation between cigarette smoking and MI. The study was performed on 789 individuals of who 267 were cigarette smokers. MI was diagnosed in 157 smokers compared to 209 non-smokers. Is there is relation between myocardial infarction and smoking? Explain your answer. o Construct a 2 x 2 table MI +ve MI -ve Total Cig Smok +ve 157 267 Cig Smok -ve 209 522 Total 789 o Calculate the incidence of myocardial infarction among smokers, nonsmokers ??I exp= 157/267= 58.8% ??I un-exp= 209/522= 40.0% o Calculate the relative risk and interpreted the findings ? RR= Iexo/Iun-exp= 58.8/40.0 ? 1.47 ? Smokers are 1.5 times more risky to develop MI compared Sir Richard Doll's study of the hazards of cigarette smoking ? One of the best examples of a prospective cohort study ??To investigate the relationship between smoking and lung cancer. ? They followed up 40 000 British doctors (divided into four cohorts): ??non-smokers, and light, moderate and heavy smokers. ? Death was the outcome they recorded. ? They used both all cause death (any death) and cause specific death (death from a particular disease). ? Results in 1964, showed a substantial excess in both mortality from lung cancer and all cause mortality in smokers, with a "dose-response" relation ?Uncontrolled trials - Experimental drug or procedure compared with another, with a placebo, or with the standard procedure - Greater validity ?Trials with independent concurrent controls ?Double or single blind ?Best is randomized assignment ?Same point in time ?These includes 4 types of controlled studies A. Randomized controlled trials (RCT): ?The epitome of all research designs ?Provides the strongest evidence of concluding causation ?Best insurance that results are due to the intervention Miss : Haneen Dhaidel , Fundemental methods research , Fall 2023 ?SINGLE-BLIND CLINICAL TRIAL: Trial in which the subject, but not the observer, does not know which of the possible treatments he is receiving.prevalence: ?In analytical cross-sectional studies, the OR can be used to assess the strength of an association between a risk factor and health outcome of interest Advantages ?Quick and cheap ?Able to measure prevalence for all factors under investigation ?Multiple outcomes and exposures can be studied ?Examine associations ?Good for generating hypotheses Disadvantages ?Snapshot in time ?Temporal associations not clear ?Selection bias ?Over-representation of cases with long duration ?Underestimate of cases with short duration.Differentially administer co-interventions ??????????????????????????????????????